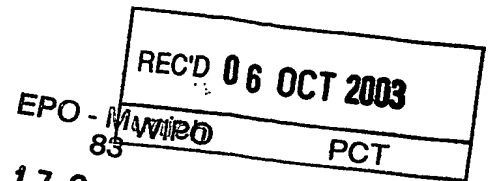


BUNDESREPUBLIK DEUTSCHLAND



EPO 3/07451

Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 102 31 613.9

Anmeldetag: 12. Juli 2002

Anmelder/Inhaber: Atvisican AG, Essen/DE

Bezeichnung: Verfahren zum Komprimieren und Dekomprimieren
von Videobildaten od. dgl.

IPC: H 04 N und G 06 T

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 9. September 2003
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Stampschloß

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Verfahren zum Komprimieren und Dekomprimieren
von Videobilddaten od. dgl.

Die Erfindung betrifft ein Verfahren zum Komprimieren und Dekomprimieren von Videobilddaten von Videobildfolgen od. dgl., die als eine Folge von jeweils in zweidimensional adressierbaren Pixeln zugehöriger Pixeldaten vorliegen, wobei jeweils die Pixeldaten ausgewählter Pixelmengen mit mathematischen Funktionen analysiert und auf deren Funktionsparamter reduziert komprimiert werden und nach einer Speicherung und/oder Übertragung mit einer entsprechend mathematischen Funktion weitgehend regeneriert dekomprimiert werden.

Derartige Verfahren sind unter den ISO-Standards MPEG, MPEG1 bis MPEG4, JPEG, etc. bekannt geworden. Bei diesen werden durch eine Differentialanalyse, Musteranalyse, Fourieranalyse od. dgl. der Pixelmengendaten von Bildausschnitten, sog. Kacheln, und insbesondere von solchen Kacheldaten in Bezug auf die Kacheldaten der Kachel mit gleichen Bildzeilen- und Bildspalten-Koordinaten von vorhergehenden Videobildern Funktionsparamter ermittelt und unter Berücksichtigung von Veränderungen in diesen Videobildfolgen verabredente Standard-Rahmenformaten gemäß dargestellt. Die Rahmenformate enthalten jeweils eine Angabe der entsprechenden Kompressionsfunktion, die jeweils umso weitergehend komprimierend ausgewählt ist je stärker der Inhalt aufeinanderfolgender Bilder oder gleichgelagerter Kacheln von solchen Bildern übereinstimmen, und die bei der Anwendung der Funktion jeweils gewonnenen Paramter.

Zur Dekompression wird jeweils aus dem Rahmen die Angabe über die jeweilige Kompressionsfunktion entnommen und dieser gemäß mittels einer entsprechenden Funktion und den bereitgestellten Parametern sowie ggf. Daten der Kachel(n) mindestens eines vorhergehenden Bildes die ursprüngliche Pixelmenge bis auf Toleranzfehler wieder hergestellt.

Es ist Aufgabe der Erfindung, eine wesentlich höhere Komprimierung der Daten im Echtzeitdurchlauf von Videobildfolgedaten bei annähernd gleicher Bildqualität wie die bekannten Verfahren zu erbringen.

Die Lösung besteht darin, daß in einer Grundanalyse der Videodaten eines Videobildes

- Konturen von Bildstrukturen anhand von sprunghaften Änderungen von Helligkeit und/oder Farbwert bei einander benachbarten Pixeln ermittelt werden,
- durch Interpolation eine Konturglättung und -schließung vorgenommen wird,
- die so gefundenen Konturen jeweils abschnittsweise durch eine parametrisierte mathematische Funktion beschrieben und als Objekte definiert werden, wobei alle Objekte, die eine Pixelanzahl unterhalb einer vorgebbaren Schwelle enthalten, einem Hintergrund zugeordnet werden,
- für die einzelnen Objekte und den Hintergrund eine Farbdominanz und ein Farbverlauf vektoriell nach Richtung und Größe ermittelt wird,
- die Lage und Erstreckung der einzelnen Objekte jeweils vektoriell bestimmt werden,
- für die einzelnen Objekte und den Hintergrund je eine Strukturfunktion ermittelt wird, und daß bei Folgeanalysen von Videobildern
- jeweils die differentiellen Veränderungen der Helligkeit, Größe, Lage und Orientierung der Objekte unter

Berücksichtigung gemeinsamer Konturen aneinanderstoßender Objekte bestimmt werden,

- die so definierten Objekte und der Hintergrund zusammen mit ihren so gewonnenen optischen, lagemäßigen und strukturellen Daten in einem strukturierten Grundrahmen oder Folgerahmen angeordnet und bereitgestellt werden;
- die demgemäß bereitgestellten Grundrahmen- und Folgerahmendaten werden zur Dekompression und einer Bildwiederaufbereitung in Pixeldaten transformiert,
- indem aus den Grundrahmendaten aus den Objekten deren zugehörigen Konturlagedaten im Pixelbild bestimmt werden,
- für den Hintergrund des Bildes und die Objekte, jeweils anhand der Konturlagedaten begrenzt, die Pixeldarstellung entsprechend der jeweilig zugehörigen Strukturfunktion mit Pixeldaten aufgefüllt werden,
- die gemäß des Farbdominanzwertes und des Farbverlaufsvektors sowie des Helligkeitwertes restituiert werden, und
- die Folgerahmendaten jeweils auf die vorherige Pixeldarstellung zur Verlagerung und/oder Veränderung der Objekte angewandt werden.

Vorteilhafte Ausgestaltungen sind in den Unteransprüchen angegeben.

Die Ermittlung und Beschreibung der Objekte anhand ihrer Konturen und deren Strukturen führt zu der extrem hohen Datenkompression im Vergleich zu den herkömmlichen Verfahren, bei denen jeweils einzelne rechteckige Ausschnitt bearbeitet werden, ohne einen größeren bildlichen Zusammenhang zu erfassen und zu nutzen.

Zur Beschleunigung des Verfahrens werden außerdem in den einzelnen Verfahrensschritten vorteilhafte neuartige Methoden angewandt, die auch als selbständige Erfindungen zu werten sind.

Aufgrund der Erkenntnis, daß sich viele Objekte in ihrer Grundstruktur und in ihrer Beziehung zu anderen gleichen, z. B. Kopf, Arme, Oberkörper, Unterkörper, Beine zu einer Person etc., werden einmal erkannte und funktionsmäßig charakterisierte Objekte anhand ihrer Daten in einem neuronalen Netz zugeordnet zu dessen weiteren und zugehörigen Objekten Konturdaten abgelegt, so daß jeweils zu einem aufgefundenen Objekt üblicher Weise angrenzende Objekte später unmittelbar aufgefunden werden und zur Erleichterung der Konturermittlung herangezogen werden können.

Auch können die Zusammenstellungen der mathematischen Funktionsbeschreibungen der verschiedenen Objekte aus dem neuronalen Netz entnommen werden, die nur mit entsprechenden aktuellen Parametern, wie Radius, Mittelpunktvektor, Anfangs- und Endkoordinaten etc. zu versehen sind.

Auch die Strukturfunktion eines Objektes ist häufig der gleichartiger Objekte gleich oder nahekommend, so daß sie als erste Annäherung dienen kann, wenn sie im Neuronennetz gespeichert ist und diesem entnommen wird.

Eine sehr hohe Kompression wird vorteilhaft durch die Nutzung der Erkenntnis erreicht, daß die Pixeldaten einer Pixelzeile jeweils eine Zahlenreihe ist, die sich durch mit natürlichen Zahlen ausgeführte elementare Rechenoperationen darstellen läßt. Insbesondere die Division und die n -te Wurzel sind einfache Operationen, die mehr oder weniger periodische Pixeldaten einer Zeile gut angenähert erbringen. Die Darstellung der Zeile schrumpft dann auf die verschlüsselte Angabe der Funktion und die Zahlenwerte, die vorzugsweise als Summe oder Differenzen von Primzahlenpotenzen dargestellt sind.

Jede bereits zu einer Pixeldatenfolge aufgefundene derartige Strukturbeschreibung wird vorzugsweise in einem neuronalen Netz abgelegt, so daß sie dort unmittelbar nutzbar oder als erste Annäherung bei späterem Vorliegen einer ähnlichen Pixeldatenfolge abzurufen ist.

Da die zu verwendenden Funktionen elementar sind und von üblichen Rechnern mit hoher Geschwindigkeit als Festkommaoperationen auszuführen sind, lassen sich aus den Strukturdaten die Pixeldaten in der Laufzeit einer Bildwiedergabe generieren; die Dekompression ist völlig problemlos.

Die Kompression von Video-Laufzeitdaten wird in ihren einzelnen Schritten vorteilhaft in ihrer Präzision an die Verträglichkeit von Abweichungen angepaßt.

Es wird bei der Konturdatenermittlung, Glättung etc. einer hohen Auflösung von bewegten Vordergrundobjekten mehr Aufmerksamkeit gewidmet als dem Hintergrund und den passiven Objekten zuteil wird, indem den Objekten jeweils zur Bearbeitung unterschiedliche maximale Rechenzeiten zugeteilt werden.

Weiterhin wird die Minimal-Anzahl der Pixel, für die ein Objekt definiert wird, jeweils einer noch verfügbaren Rechenzeit angepaßt. Die größten Objekte werden zuerst bearbeitet und soweit noch Rechenzeit für eine Bildzeit übrig ist, werden kleinere Objekte aus dem Hintergrund herausgetrennt und detailliert geometrisch und strukturell beschrieben und in den Rahmen eingestellt.

Für die Ermittlung einer Strukturfunktion eines Objektes wird vorteilhaft jeweils eine Maximalzeitvorgabe vorgenommen, wobei die Erkenntnis genutzt wird, daß Abweichungen der einzelnen Pixeldaten, wenn sie nicht benachbart gehäuft auftreten, keine

wesentlichen Bildqualitätseinbußen erbringen, da die Struktur nur das allgemeine Aussehen der Oberfläche eines Objektes aber keine Bilddetails betrifft.

Als Beispiel einer Strukturfunktion sei zur Veranschaulichung genannt:

x -te Wurzel aus a hoch m \pm b hoch n durch c hoch p \pm d hoch q ; x = ganzzahlig $1 \div 3$; a, b, c, d = Primzahlen bis 17; m, n, p, q = ganzzahlig $1 \div 9$.

Als zu analysierende Pixelmenge werden beispielsweise jeweils 256 Pixel eines Bildzeilenabschnittes oder eines 8×8 oder 16×16 Pixelbild-Ausschnittes genommen. Die Pixeldaten sind gewöhnlich in 8-bit verschlüsselt. Demgemäß werden die Operationen nicht dezimal oder hexadezimal sondern modulo 256 ausgeführt, so daß die Quelldaten wie die Verschlüsselungsdaten und die zurückgewonnenen Zieldaten stets als 8-bit Pixeldaten unmittelbar vorliegen.

Werden mehrere Zeilenabschnitte einer Bildzeile oder aufeinanderfolgender Bildzeilen analysiert, so ergibt sich häufig aus einer Fortführung und/oder aus einer Versetzung um einige Stellen der zuvor zutreffenden Strukturfunktion eine passende Lösung in sehr einfacher und zeitsparender Weise. Statt einer neuen Strukturfunktion wird dann die Abwandlung im zugehörigen Rahmen angegeben.

Fig. 1 zeigt ein Blockschaltbild der Bildverschlüsselung.

Die Videodaten VD werden nach und nach den verschiedenen Verfahrensschritten unterzogen.

Als erstes erfolgt die Objekterkennung OE, wobei die im Bild zuvor erkannten Objekte 01*; 02* sowie die in einem ersten

neuronalen Netz NN1 abgelegten Objekte als Hilfsinformationen genutzt werden. Die erkannten Objekte werden einer Objektglättung OG unterzogen, wobei eine Auflösungsgrenze MIN vorgegeben ist.

Die geglätteten Objekte werden einer Objektbeschreibung unter Beachtung der Nachbarschaftsgrenzbeziehungen unterzogen, so daß die Objekte O1, O2 etc. funktional in dem Rahmen FR abgelegt werden.

Zu den einzelnen Objekten erfolgt die Festlegung OLV der Lage- und Richtungsvektoren OL1, OL2 etc. sowie die Farbbeschreibung OFV mittels der Farb- und Farbverlaufsvektoren OF1, OF2 etc.

Weiterhin werden zu den Objekten O1, O2 etc. die Strukturfunktionen und deren Parameter OS1, OS2 etc. vorzugsweise unter Zuhilfenahme eines zweiten neuronalen Netzes NN2 bestimmt und im Rahmen FR abgelegt, ebenso wie die Lage- und Farbvektoren.

Sind alle Objekte im Rahmen festgehalten, werden vom Hintergrund HG noch die Farbvektoren HGF und die Hintergrundstrukturen HGS ermittelt und in den Rahmen FR abgestellt. Ein kompletter Rahmen FR eines Bildes wird dann als historischer Rahmen FRH bereitgestellt, dessen Inhalte, die jeweils mit einem Stern am Referenzzeichen gekennzeichnet sind, der Verschlüsselung des nächsten Bildes als Ausgangsmaterial zur Verfügung gestellt.

Falls nur geringe Änderungen der Farbe, Lage, Struktur oder Orientierung eines Objektes festgestellt wird, werden nur die Änderungen im Folgerahmen spezifiziert, was eine erhebliche Einsparung an Verarbeitungszeit, Speicher- und Übertragungskapazität erbringt.

Jeweils aufgefundene Objektbeschreibungen, ihre Nachbarschaftsbeziehungen sowie die Strukturfunktionen werden den Basen der neuronalen Netze NN1, NN2 zugeführt, so daß ähnliche Objekte und Strukturen aufgefunden und bei der Verschlüsselung neuer Bilder genutzt werden.

Die Verschlüsselungszeit wird jeweils über einen Zeitmanager TMG überwacht und durch geeignete Vorgaben der Mindestauflösung MIN und der Maximalzeit TMAX der Strukturanalyse in Grenzen gehalten.

Eine Alternative der Berechnung der Strukturfunktionen zur zuvor beschriebenen läßt sich ähnlich vorteilhaft mit hexadezimalen Operationen vornehmen, wozu die üblichen 8-bit-Pixelinformationen in zwei 4-bit-Zeichen geteilt werden und so die doppelte Stellenzahl berechnet und auf größtmögliche Ähnlichkeit geprüft wird. Die Funktionen und deren Parameter werden zweckmäßig, insbesondere in dem Zusammenhang, auch als Hexadezimalziffern verschlüsselt und paarweise in 8-bit Bytes im Rahmen verpackt. Je nach der angegebenen Funktion sind mehr oder weniger Parameter anzugeben.

Eine sehr hohe Packungsdichte im Rahmen läßt sich auch dadurch erreichen, wenn in einem Byte jeweils drei Bit für acht Funktionen, 3 Bit für die acht ersten Primzahlen und zwei Bit für deren Exponenten von 1 - 4 abgelegt werden. Als Funktionselemente lassen sich z. B. die vier Grundrechenarten, die Wurzel- und Potenzfunktion sowie Formelklammern verschlüsseln. Zu den Klammerfunktionen lassen sich in den weiteren 5 bit des Bytes weitere Sonderfunktionen angeben, wie ein Formelendezeichen oder komplexe Funktionen.

Patentansprüche

1. Verfahren zum Komprimieren und Dekomprimieren von Videobilddaten von Videobildfolgen od. dgl., die als eine Folge von jeweils in zweidimensional adressierbaren Pixeln zugehöriger Pixeldaten vorliegen, wobei jeweils die Pixeldaten ausgewählter Pixelmengen mit mathematischen Funktionen analysiert und auf deren Funktionsparameter reduziert komprimiert werden und nach einer Speicherung und/oder Übertragung mit einer entsprechenden mathematischen Funktion weitgehend regeneriert dekomprimiert werden, dadurch gekennzeichnet, daß in einer Grundanalyse der Videodaten eines Videobildes
 - Konturen von Bildstrukturen anhand von sprunghaften Änderungen von Helligkeit und/oder Farbwert bei einander benachbarten Pixeln ermittelt werden,
 - durch Interpolation eine Konturglättung und -schließung vorgenommen wird,
 - die so gefundenen Konturen jeweils abschnittsweise durch eine parametrisierte mathematische Funktion beschrieben und als Objekte definiert werden, wobei alle Objekte, die eine Pixelanzahl unterhalb einer vorgebbaren Schwelle enthalten, einem Hintergrund zugeordnet werden,
 - für die einzelnen Objekte und den Hintergrund eine Farbdominanz und ein Farbverlauf vektoriell nach Richtung und Größe ermittelt wird,
 - die Lage und Erstreckung der einzelnen Objekte jeweils vektoriell bestimmt werden,
 - für die einzelnen Objekte und den Hintergrund je eine Strukturfunktion ermittelt wird, und daß bei Folgeanalysen von Videobildern
 - jeweils die differentiellen Veränderungen der Helligkeit, Größe, Lage und Orientierung der Objekte unter Berücksichtigung gemeinsamer Konturen aneinanderstoßender Objekte bestimmt werden,

- die so definierten Objekte und der Hintergrund zusammen mit ihren so gewonnenen optischen, lagemäßigen und strukturellen Daten in einem strukturierten Grundrahmen oder Folgerahmen angeordnet und bereitgestellt werden;
- die demgemäß bereitgestellten Grundrahmen- und Folgerahmendaten werden zur Dekompression und einer Bildwiederaufbereitung in Pixeldaten transformiert, indem aus den Grundrahmendaten aus den Objekten deren zugehörigen Konturlagedaten im Pixelbild bestimmt werden,
- für den Hintergrund des Bildes und die Objekte, jeweils anhand der Konturlagedaten begrenzt, die Pixeldarstellung entsprechend der jeweilig zugehörigen Strukturfunktion mit Pixeldaten aufgefüllt werden,
- die gemäß des Farbdominanzwertes und des Farbverlaufsvektors sowie des Helligkeitswertes restituiert werden, und
- die Folgerahmendaten jeweils auf die vorherige Pixeldarstellung zur Verlagerung und/oder Veränderung der Objekte angewandt werden.

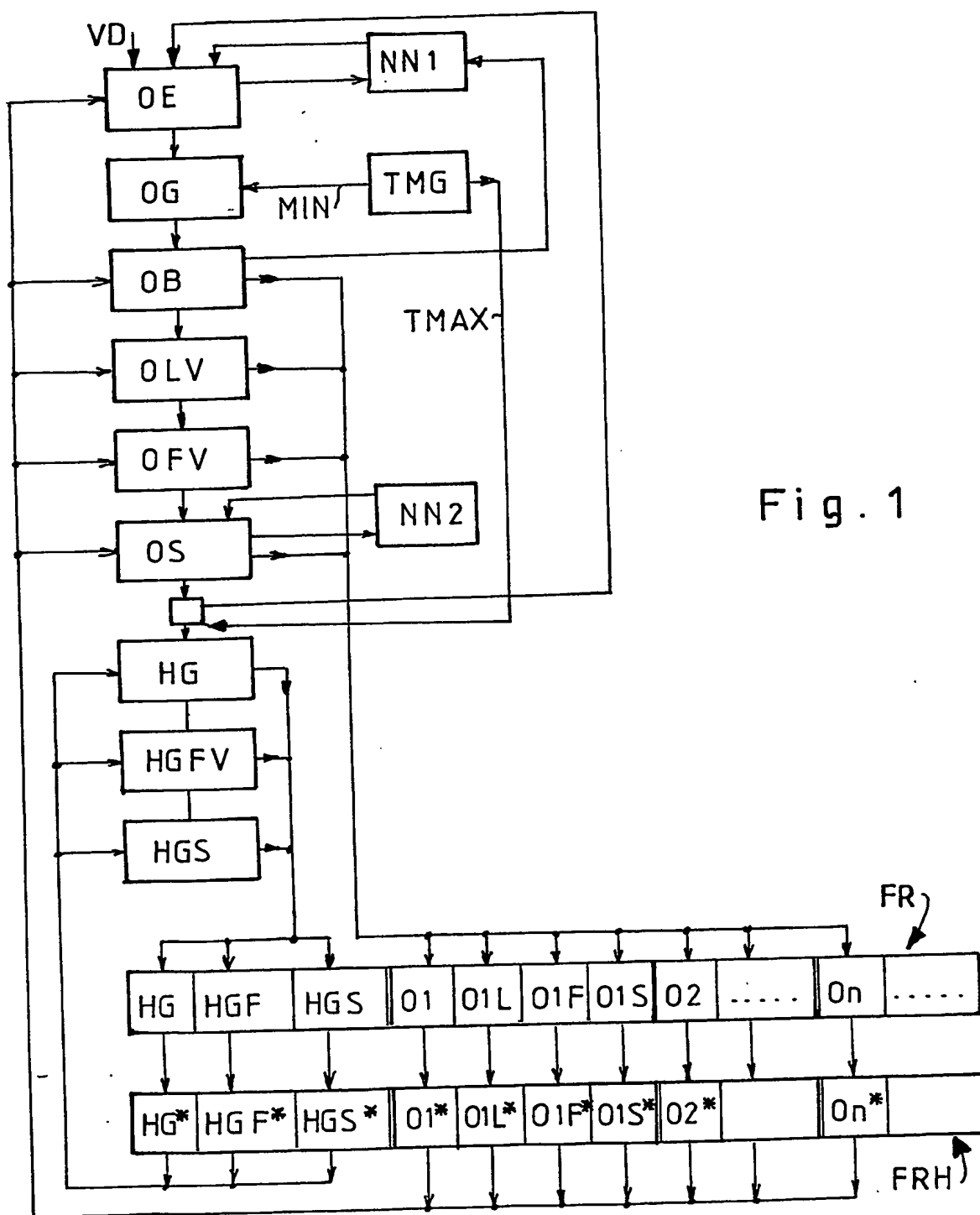
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die beschriebenen Objekte mit ihren mathematischen Funktionen in einem neuronalen Netz (NN1) abgelegt werden, das der weiteren Erkennung (OE) von Objekten in Videobilddaten (VD) dient.
3. Verfahren nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß ermittelte Strukturfunktionen (OS) mit deren Parametern von Objekten und Hintergründen in einem neuronalen Netz (NN2) gespeichert werden, das bei der weiteren Ermittlung von Strukturfunktionen (OS) mit deren Parametern als eine Ausgangsbasis dient.
4. Verfahren nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß die Strukturfunktion (OS) jeweils als

eine mathematische Funktion dargestellt ist und die Parameter ganzzahlige Größen sind und die Funktion eine unbegrenzte Zahl von Nachkommastellen liefert.

5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß die Strukturfunktion (OS) ein Bruch, eine n-te Wurzel oder eine transzendente Funktion ist.
6. Verfahren nach Anspruch 4 oder 5, dadurch gekennzeichnet, daß die ganzzahligen Größen als Potenzen von Primzahlen sowie Summen oder Differenz von diesen verschlüsselt dargestellt sind.
7. Verfahren nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet, daß die Parameter modulo 2 hoch 8 dargestellt sind und die Funktion mit Größen, die modulo 2 hoch 8 dargestellt sind, ausgeführt werden und solche Größen als Nachkommastellen liefern.
8. Verfahren nach einem der Ansprüche 4 bis 7, dadurch gekennzeichnet, daß die einzelnen Strukturfunktionen (OS) jeweils annähernd passend zu einer Pixeldatenfolge eines Bildzeilenabschnittes einer vorgegebenen Länge oder eines rechteckigen Pixelbildausschnittes bestimmt werden.
9. Verfahren nach Anspruch 8, dadurch gekennzeichnet, daß der Zeilenabschnitt eine Länge von 64, 128 oder 256 Bytes hat oder der Pixelbildausschnitt eine Größe von 8 mal 8 oder 16 mal 16 Bytes aufweist.
10. Verfahren nach einem der Ansprüche 8 oder 9, dadurch gekennzeichnet, daß die Strukturfunktion (OS) jeweils solange oder so genau durch sukzessive Approximation an die jeweils annähernd darzustellende Pixeldatenfolge angepaßt wird, die durch eine Zeitvorgabe (TMax) oder eine

Genauigkeitsvorgabe bestimmt ist.

11. Verfahren nach Anspruch 10, dadurch gekennzeichnet, daß die Zeit-oder Genauigkeitsvorgabe abhängig von der Lage oder einer jeweiligen Lageveränderungsgeschwindigkeit des jeweiligen Objektes bestimmt wird, wobei für mittig im Bild liegende und/oder ruhende Objekte eine längere Zeit und/oder eine höhere Genauigkeit zugeteilt wird als randseitigen und/oder relativ schnell bewegten Objekten und/oder dem Hintergrund.
12. Verfahren nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß jeweils nur solche Objekte einer weiteren Erkennung und Charakterisierung unterzogen werden, die eine Mindestanzahl von Pixeln aufweist, und kleinere Objekte dem Hintergrund zugeordnet werden.
13. Verfahren nach Anspruch 12, dadurch gekennzeichnet, daß die Objekte mit fallender Anzahl von Pixeln nacheinander bearbeitet werden solange die verfügbare Rechenzeit reicht, wodurch bei der Verschlüsselung eines Bildinhaltes die Mindestanzahl der Pixel der Objekte je nach der verfügbaren Rechenzeit bestimmt ist.



Zusammenfassung

Verfahren zum Komprimieren und Dekomprimieren
von Videobilddaten od..dgl.

Verfahren zum Komprimieren und Dekomprimieren von Videobilddaten, wobei in einer Grundanalyse der Videodaten eines Videobildes Konturen von Bildstrukturen anhand von sprunghaften Änderungen von Helligkeit und/oder Farbwert bei einander benachbarten Pixeln ermittelt werden, die so gefundenen Konturen jeweils abschnittsweise durch eine parametrisierte mathematische Funktion beschrieben und als Objekte definiert werden, für die einzelnen Objekte eine Farbdominanz und ein Farbverlauf sowie die Lage und eine Erstreckung der einzelnen Objekte und eine Strukturfunktion ermittelt wird, und daß bei Folgeanalysen von Videobildern jeweils die differentiellen Veränderungen der Helligkeit, Größe, Lage und Orientierung der Objekte unter Berücksichtigung gemeinsamer Konturen aneinanderstoßender Objekte bestimmt werden. Die so definierten Objekte werden in einem strukturierten Grundrahmen oder Folgerahmen angeordnet und bereitgestellt. Die Konturanalyse und die Strukturanalyse wird mittels neuronaler Netze vorgenommen.

Fig. 1

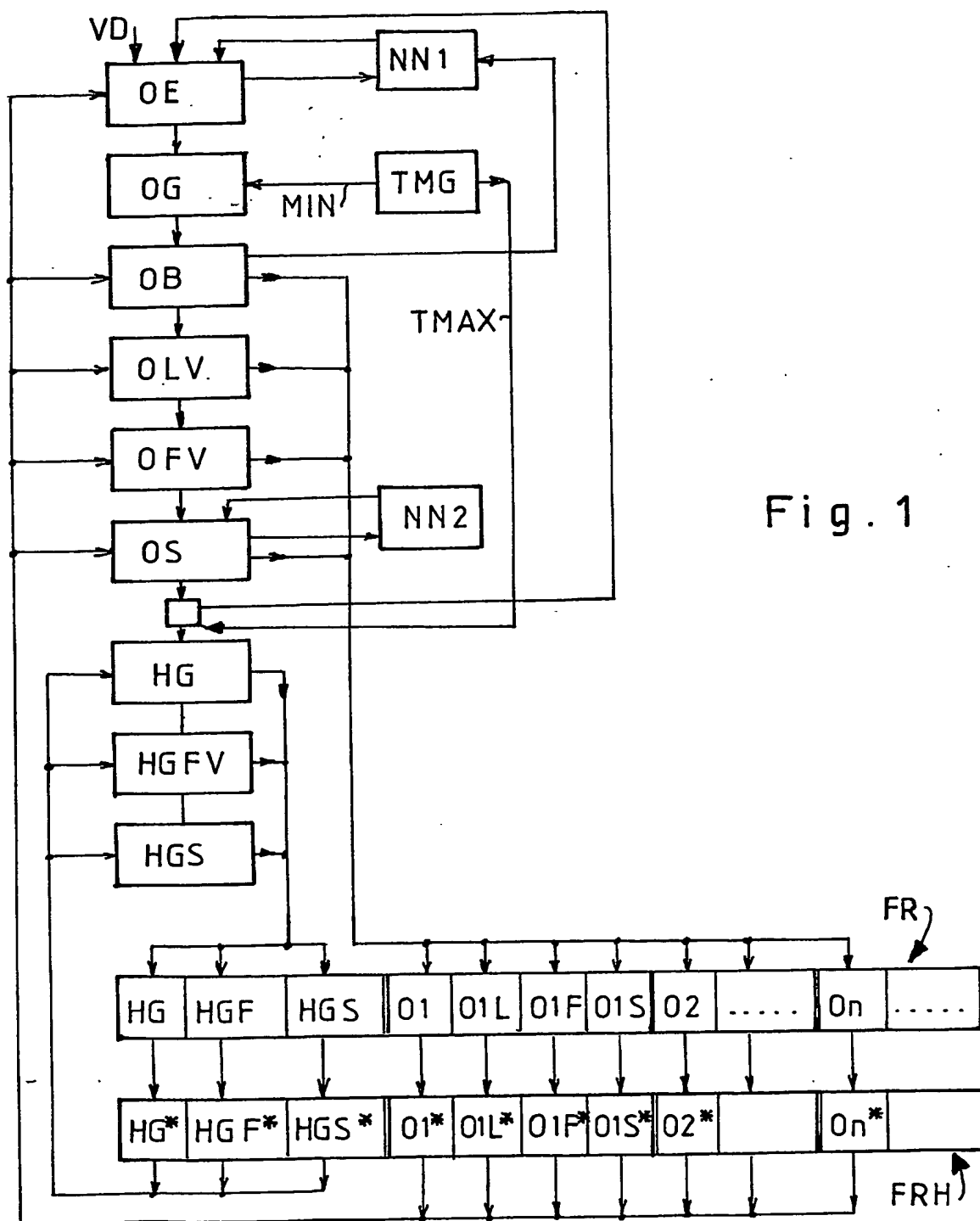


Fig. 1